Light Field Image Denoising using a Linear 4D Frequency-Hyperfan All-In-Focus Filter

Donald G. Dansereau, Daniel L. Bongiorno, Oscar Pizarro, and Stefan B. Williams

Australian Centre for Field Robotics; School of Aerospace, Mechanical and Mechatronic Engineering; University of Sydney, NSW, Australia

ABSTRACT

Imaging in low light is problematic as sensor noise can dominate imagery, and increasing illumination or aperture size is not always effective or practical. Computational photography offers a promising solution in the form of the light field camera, which by capturing redundant information offers an opportunity for elegant noise rejection. We show that the light field of a Lambertian scene has a 4D hyperfan-shaped frequency-domain region of support at the intersection of a dual-fan and a hypercone. By designing and implementing a filter with appropriately shaped passband we accomplish denoising with a single all-in-focus linear filter. Drawing examples from the Stanford Light Field Archive and images captured using a commercially available lenselet-based plenoptic camera, we demonstrate that the hyperfan outperforms competing methods including synthetic focus, fan-shaped antialiasing filters, and a range of modern nonlinear image and video denoising techniques. We show the hyperfan preserves depth of field, making it a single-step all-in-focus denoising filter suitable for general-purpose light field rendering. We include results for different noise types and levels, over a variety of metrics, and in real-world scenarios. Finally, we show that the hyperfan's performance scales with aperture count.

1. INTRODUCTION

Many vision applications call for operation in light- or contrast- limited environments, including operation at night, underwater, in smoke or fog, or with short exposure times. In any scenario for which illumination or contrast is limited, sensor noise can dominate imagery and impede visual processing. It is not always possible to ameliorate imagery by increasing illumination: large scenes cannot always be effectively lit, illumination power budgets can be limited, and in turbid (murky) water or foggy air, backscatter can negate any advantage gained by increasing illumination.

Given that increasing illumination is not always practical, modifying the camera to gather more light seems like the next logical approach. Unfortunately, the straightforward steps of increasing aperture size or exposure duration have the evident drawbacks of reducing depth of field or increasing motion blur, respectively. Meanwhile, computational photography is gaining prominence for its ability to offer powerful computational approaches to traditionally physical problems. Light field cameras, in particular, are emerging as a powerful tool for capturing rich scene representations that offer post-capture refocus, increased effective aperture size, and increased depth of field.[1–5]

We propose the use of light field cameras in difficult imaging scenarios, exploiting the redundant information they capture to reject unwanted noise while preserving the desired signal. Working from the properties of parallax motion within the light field, we describe a limited subset of the frequency domain, the 4D hyperfan, in which such images must exist. This allows us to design and implement a filter with a passband which effectively surrounds this frequency-domain region of support (ROS), passing the light field signal while rejecting a significant subset of the noise.

We provide background on light fields and the problem of denoising in Section 2. Section 3 discusses the light field characteristics that are exploited in Section 4 to derive the filter. Section 5 shows results typical of the filter for real-world light fields, and quantifies its performance. Finally, Section 6 draws conclusions and indicates directions for future work.

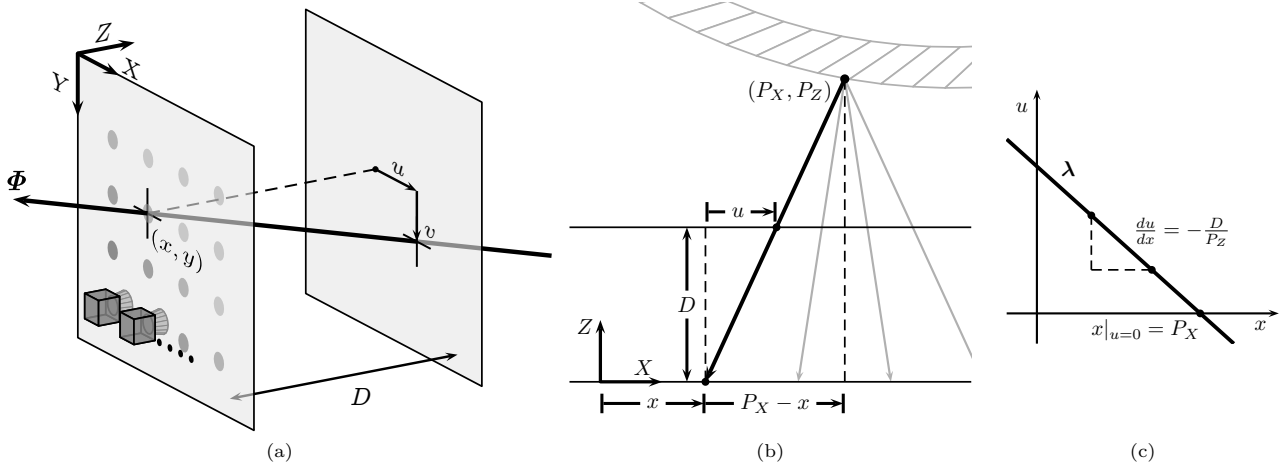Author email addresses: {d.dansereau, d.bongiorno, o.pizarro, s.williams}@acfr.usyd.edu.au

Figure 1. (a) The local two-plane parameterization of light rays; (b) for all rays originating at a point P in space, u varies linearly with x, and by extension v with y; (c) this describes a plane λ in the 4D light field, shown here in the (x, u) plane

2. BACKGROUND

Denoising of images is a well-explored topic, and a good review is provided by Wilburn et al.[6] See also Ref. 7 for modern overcomplete dictionary developments, and Refs. 8 and 9 for a singular value decomposition generalization of K-means for learning dictionaries directly from noisy imagery. Because we are dealing with high-dimensional imagery, video denoising is also relevant, including recent advances in block matching and filtering.[10]

Alternative approaches to low-light and contrast-limited imaging have recently appeared in the realm of computational photography. Levoy et al.[11] demonstrate an active illumination generalization of confocal imaging, allowing effective imaging through turbid media. Turning that idea on its head by varying the position of the camera rather than the illumination source yields the light field-based approach explored in the present work. A similar idea is explored by Bishop and Favaro,[1] employing a variational Bayesian framework to combine light measured across many apertures. Our work differs significantly in its complexity: we present a single linear filter as a means of combining images from across the light field, offering a far simpler and potentially more robust solution.

In other relevant work a dual-fan shaped passband is utilized for linearly filtering light fields.[12] Although its 2D separability makes efficient implementation of the dual-fan possible, we demonstrate that its passband is suboptimal, being significantly outperformed by the more selective hyperfan.

2.1 The Light Field

Whereas a conventional camera measures variations in light as a function of direction for rays passing through a single position (aperture), a light field camera encodes variations in light as a function of both direction and position. This is achieved by placing coded masks[2] or lens arrays[3] in the optical path of a conventional camera, or by building a grid of conventional cameras[5] as depicted in Fig. 1a. The resulting images open a range of new processing possibilities, including featureless or linear methods for traditionally nonlinear, iterative or nondeterministic tasks. A few examples include focus and depth-selective filtering, distractor isolation and visual odometry.[3, 12–15]

A convenient way to parameterize the light field is using the four-dimensional local two-plane parameterization (L2PP)[15] depicted in Fig. 1a, in which each of the measured rays is described by its points of intersection with two reference planes: the (x, y) plane given by $z = 0$, and a "local" (x, y)-centered (u, v) plane which is parallel to the (x, y) plane at a positive separation $z = D$. A coordinate in this space describes a ray $\boldsymbol{\Phi} = (x, y, u, v)$, where (x, y) defines the *position* of the ray, and (u, v) defines its *direction*. A light field encoded in the original two-plane parameterization (2PP)[16] contains the same information, and the filter we propose operates with either representation via a straightforward modification of filter parameters.

Similarly, because all 4D light field cameras measure the same form of information, the filters described in the following sections can operate on images measured by any model, requiring at most a reparameterization of the light field or modification of filter parameters. In terms of suitability for low-contrast imaging, we note that an array of conventional cameras offers the unique advantage of scalability: each camera in the array has its own aperture and sensor area, and so adding cameras results in more light being gathered. We also note that lenselet-based light field cameras offer significantly wider apertures for a given depth of field than conventional cameras, with the advantage being proportional to the number of pixels beneath each lenselet.[1,3,4] Mask-based cameras attenuate incoming light by virtue of their modulating masks, potentially making them less suitable for low-contrast imaging than other models.

2.2 Point-Plane Correspondence and Frequency Planes

As depicted in Fig. 1b, all rays emanating from a point $\boldsymbol{P}$ in a scene are constrained by two simultaneous linear equations

$$\begin{bmatrix} u \\ v \end{bmatrix} = \left(\tfrac{D}{P_Z}\right) \begin{bmatrix} P_X - x \\ P_Y - y \end{bmatrix}, \tag{1}$$

where each equation is similar to parallax motion in a moving camera.[17,18] The linear relationship in (x, u) describes a *hyperplane* in 4D space, and by extension a similar hyperplane exists in (y, v). Taken simultaneously these hyperplanes intersect to describe the plane $\boldsymbol{\lambda}$, a slice of which is depicted in (x, u) in Fig. 1c.

In a Lambertian scene free of occlusions the observed values emanating from a surface element at $\boldsymbol{P} = (P_X, P_Y, P_Z)$ will be constant with angle, and so the values on the corresponding light field plane $\boldsymbol{\lambda}$ will also be of constant value.[18] A scene consisting of surface elements at a single depth will therefore yield a light field of parallel, constant-valued planes with orientation determined by the surface's depth, following (1). The Fourier transform of such a set of parallel planes is a single plane through the origin[18] – this arises because the light field is constant-valued in two directions parallel to the constant-valued spatial planes. The frequency plane can be described as

$$\Omega_x/\Omega_u = \Omega_y/\Omega_v = D/P_Z - 1. \tag{2}$$

Generalizing for a scene containing a range of depths is possible through superposition: a scene comprising surface elements at many depths will exist as a superposition of planes in the light field. This can be seen by allowing P_Z in (2) to sweep through a range of depths corresponding to the scene extents,

$$Z_{MIN} < P_Z < Z_{MAX}. \tag{3}$$

The nature of the resulting shape is explored in the following section.

We explicitly ignore the effects of occlusion, for which planes in the light field are truncated, and non-Lambertian surfaces, for which rays within the plane have different values. Many practical scenes have relatively little energy in these components, and we shall demonstrate that the filters we derive are effective despite their presence.

3. THE LIGHT FIELD AS A FREQUENCY HYPERFAN

In this section we derive the frequency-domain ROS of a light field, arriving at the conclusion that it is a hyperfan lying at the intersection of a hypercone and a dual-fan. Taking (2) and (3) together, we can write three constraints to describe the frequency-domain ROS of the light field:

$$m_{MIN} < \Omega_x/\Omega_u < m_{MAX}, \tag{4}$$

$$m_{MIN} < \Omega_y/\Omega_v < m_{MAX}, \tag{5}$$

$$\Omega_x/\Omega_u = \Omega_y/\Omega_v. \tag{6}$$

The first two constraints (4), (5) describe 2D fan shapes in (Ω_x, Ω_u) and (Ω_y, Ω_v), respectively. The simultaneous combination of these two constitutes a dual-fan.[12] The third constraint (6), described by Levin et al.[19] in the
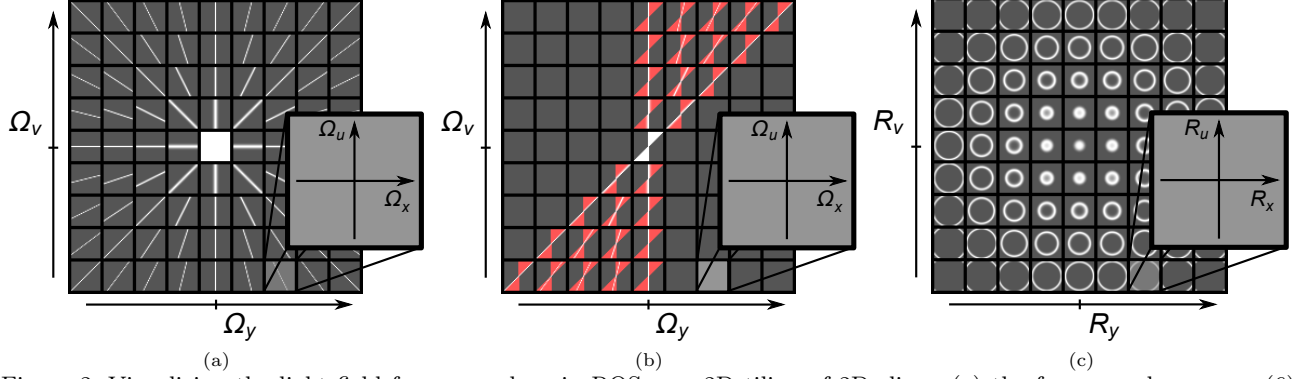
Figure 2. Visualizing the light field frequency-domain ROS as a 2D tiling of 2D slices; (a) the frequency-hypercone (6) constrains slopes in two pairs of dimensions; (b) the dual-fan (4), (5), shown in red, lacks the hypercone constraint; the fully constrained hyperfan, shown in white, is at the intersection of the hypercone and the dual-fan; (c) the rotated hypercone more clearly shows its conic nature

context of lens design, forces a plane's slope in (Ω_x, Ω_u) to equal its slope in (Ω_y, Ω_v). By combining these constraints we will formulate a filter which is significantly more selective than those previously described.

The third constraint alone, which we shall demonstrate describes a *hypercone*, is depicted in Fig. 2a as a set of (Ω_x, Ω_u) tiles in (Ω_y, Ω_v). A typical dual-fan ROS is shown in Fig. 2b in red, while the more selective ROS at the intersection of the dual-fan and the hypercone is shown in white. Because this new ROS can be formed by sweeping a plane through 4D space, we denote it a *hyperfan*.

To see why (6) describes a hypercone, we begin with the standard form

$$R_x^2 + R_u^2 - R_y^2 - R_v^2 = 0, \tag{7}$$

which describes a 4D saddle or hyperbolic cone – this differs from the 4D spherical cone in the sign of the third term. To show equivalence with (6), we transform the coordinate axes by applying rotations of $-\pi/4$ in the (Ω_x, Ω_v) and (Ω_y, Ω_u) planes, yielding

$$\begin{bmatrix} R_x \\ R_y \\ R_u \\ R_v \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \Omega_x + \Omega_v \\ \Omega_y + \Omega_u \\ \Omega_y - \Omega_u \\ \Omega_x - \Omega_v \end{bmatrix}. \tag{8}$$

Substituting the rotated coordinates into (7) and simplifying yields the form shown in (6), thus the two forms are rotated views of the same shape. The rotated form of the hypercone (7) is depicted in Fig. 2c.

4. THE 4D HYPERFAN FILTER

Having described a frequency-domain ROS for the light field, all that remains is to design a filter capable of selectively passing it. In this initial exploration we describe the filter directly in the frequency domain, implementing it by computing the discrete Fourier transform (DFT), multiplying by the filter's magnitude response in the frequency domain, and then computing the inverse DFT.

Because the frequency hyperfan lies at the intersection of a dual-fan and a hypercone as depicted in Fig. 2, one way forward is to describe each of those passbands and take their product. As we proceed we will evaluate the theoretic selectivity of each passband as the fractional 4D Nyquist hypervolume that it passes, with smaller fractions corresponding to higher selectivity.

Starting with the dual-fan passband, we note that this is itself the product of two 2D fan filters[12]

$$H_{DF}(\boldsymbol{\Omega}) = H_{FAN}^{2D}(\Omega_x, \Omega_u, \theta_{MIN}, \theta_{MAX}) \, H_{FAN}^{2D}(\Omega_y, \Omega_v, \theta_{MIN}, \theta_{MAX}), \tag{9}$$

where each 2D fan is straightforwardly implemented by passing all points within the prescribed angle range θ_{MIN} to θ_{MAX}. The 2D fan filter is explored by Ansari,[20] and the process of selecting θ values for a desired depth range is described in Ref. 12 – note that this process will vary with light field geometry and parameterization.

The fractional 2D area passed by each 2D fan has a lower bound α_{DF} determined by the range $[\theta_{MIN}, \theta_{MAX}]$. We apply Gaussian smoothing to reduce ringing artifacts, surrounding the fan by a tunable bandwidth and increasing the passband area by β_{DF}. Because the same selectivity is applied in (Ω_x, Ω_u) and in (Ω_y, Ω_v), the fractional volume passed by the 4D dual-fan is given by the square

$$V_{DF} = (\alpha_{DF} + \beta_{DF})^2. \tag{10}$$

The ideal hypercone (6) is a 3D manifold, not a 4D volume, and so practical implementation requires surrounding the hypercone by a bandwidth β_{HC}. We propose the filter with magnitude response

$$H_{HC}(\boldsymbol{\Omega}) = \exp\left(-\sqrt{2\ln 2}\left[\frac{(\Omega_x\Omega_v - \Omega_y\Omega_u)}{\beta_{HC}^2}\right]^2\right), \tag{11}$$

where β_{HC} is the 3-dB bandwidth measured as the radius of the hypercone at the origin – this is the radius of the cone in the rotated (R_x, R_u) or (R_y, R_v) plane. The magnitude of the numerator of the exponential increases with distance from the ideal hypercone shape, and so the filter rolls off in a Gaussian-like manner from the ideal passband. Note that the filter offers no selectivity near the origin, but this is consistent given that the underlying constraint (6) provides no information to do so.

For analysis we begin by ignoring the Gaussian rolloff, approximating the hypercone filter as having constant thickness related to the 3-dB bandwidth β_{HC} through a constant factor κ. Examining Fig. 2a, this implies every (Ω_x, Ω_u) slice, with the exception of the origin, will pass a constant fraction of its area. Including the effect of the Gaussian rolloff increases the total admitted volume by another constant factor which we absorb into κ, for a fractional volume passed by the hypercone given by

$$V_{HC} = \kappa_{Gauss}\beta_{HC}. \tag{12}$$

The hyperfan filter is simply the product of the hypercone and dual-fan

$$H_{HF} = H_{HC}H_{DF}. \tag{13}$$

Referring to Fig. 2b, we notice that every nonzero (Ω_x, Ω_u) slice of the hyperfan will pass a mean area of $\kappa_{Gauss}\beta_{HC}$, and that $\alpha_{DF} + \beta_{DF}$ describes the ratio of nonzero slices. The fractional volume passed by the hyperfan filter is therefore the product

$$V_{HF} = \kappa_{Gauss}\beta_{HC}(\alpha_{DF} + \beta_{DF}). \tag{14}$$

Notice the minimum volume passed by the dual-fan is α_{DF}^2, while the minimum for the hyperfan is zero – i.e. the hyperfan offers direct control, via β_{HC}, of the total signal energy passed, and therefore presents significantly greater selectivity than the equivalent dual-fan filter. Note also that both the dual-fan and hyperfan filters degenerate gracefully to frequency-planar filters as their depth ranges approach zero.

5. RESULTS

5.1 The Data

The Stanford Light Field Archive* is a publicly accessible database suitable for evaluating light field filtering techniques. The twelve light fields we utilize, listed in Table 1, all contain 17×17 aperture positions (x, y). Aperture positions are close enough to an ideal grid that ignoring the deviation results in negligible degradation to output quality. Each image in (x, y) is rectified in (u, v), and the light fields are in the 2PP parameterization. Light field geometry varies across the data set: grid spacing is not identical, plane separation varies, and image

---
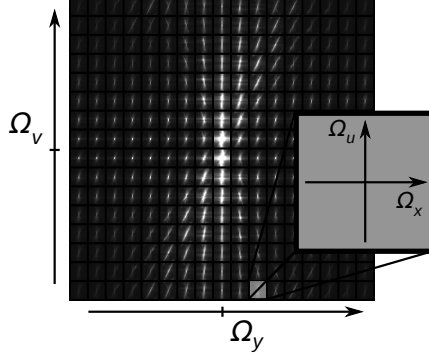
*http://lightfield.stanford.edu/

Figure 3. The maximum magnitude per frequency component over the first six Stanford light fields, showing a hyperfan shape (compare with Fig. 2b)

aspect and resolution vary, meaning fan extents θ need to be tuned on a per-light field basis. An alternative would have been to convert the light fields to a uniform L2PP parameterization and use generic fan bounds.

For all results, the Stanford light fields are downsampled to a maximum (u, v) size of 128×128 pixels. With the exception of displayed colour images, the results are for monochrome versions of the light fields. When an experiment calls for less than 17×17 apertures we discard apertures at the edge of the light field, retaining the central portion – this emulates a camera array, rather than a lens-array or mask-based camera. For consistency across experiments for which aperture counts can vary, metrics report on the central image in (x, y).

As empirical evidence of the frequency-hyperfan ROS of light fields, we computed the DFT of the first six of the twelve Stanford light fields (as listed in Table 1), scaled to a common size, and selected the maximum magnitude at each frequency. The result, shown in Fig. 3, establishes the bounds of the light fields in frequency space: the hyperfan shape is clearly evident. Note that this is true despite the varying light field geometries and the presence of occlusions and non-Lambertian surfaces.

Further qualitative validation of the hyperfan filter is carried out on imagery collected using a commercially available Lytro lenselet-based light field camera. This imagery includes low-light and turbid media examples. The raw lenselet images are decoded to a 9×9 array of images, each having 380×380 pixels. Image rectification is not carried out, and is left as future work. Compared with the 17×17 images of the Stanford light fields, we expect significantly less selectivity. However, there is still a potential 81-fold redundancy in the imagery (actually slightly less due to lenselet vignetting) and so significant noise rejection is possible.

5.2 The Methods

We test a range of linear filters on the Stanford light fields, including the three described in this paper: the hyperfan (13), the hypercone (11) and the dual-fan (9). If our earlier assertions are correct, the hyperfan will be the most selective of these, though how the hypercone alone behaves will also prove interesting.

We further test a 4D Gaussian filter as well as a 4D planar Gaussian filter which is the basis for synthetic refocusing of light fields.[3, 4, 18] Dictionary-based image denoising approaches do not exploit the structure of the light field, nevertheless by collapsing the light field into a tiling of images we test the overcomplete discrete cosine transform (DCT)[7] and K-SVD methods.[8, 9] Finally, we test the block-matching and filtering approach V-BM3D[10] by applying it over sequences along the x dimension.

5.3 Tuning

The hyperfan has four tunable parameters: the two depth limits and filter rolloff associated with the dual-fan filter, and the bandwidth associated with the hypercone. The optimal values for these depend on the range of depths occupied by the scene, the number of apertures in the light field, the noise level, and the light field parameterization.

If no prior knowledge of scene depth is available, a great deal of selectivity is nevertheless possible, as the valid range of plane angles present in any light field is limited.[19] In the L2PP, for example, all planes must lie
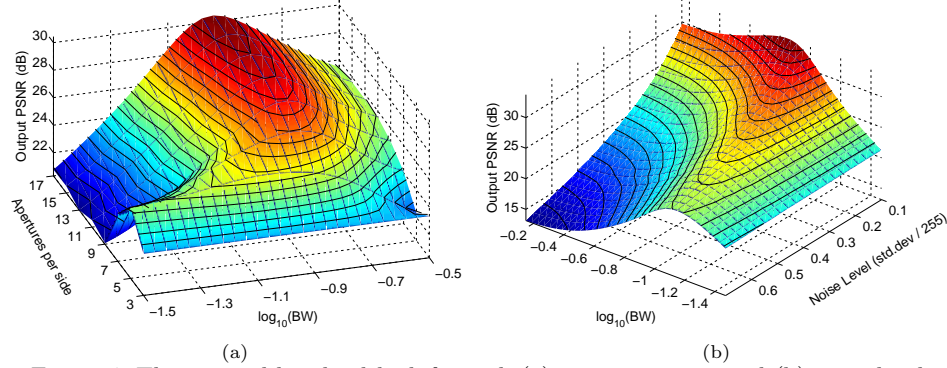
Figure 4. The optimal bandwidth shifts with (a) aperture count and (b) noise level

within the first and third quadrants in (Ω_x, Ω_u) and (Ω_y, Ω_v) – i.e. the plane angles are restricted to a ninety degree range. This observation allows the fan limits to be pre-tuned for generic scenes, leaving only the hypercone bandwidth to be tuned. Of course, knowledge of a more selective depth range allows for more aggressive filtering.

For fixed fan angles and selectivity, Fig. 4 demonstrates the dependence of the optimal hypercone bandwidth on input noise level and aperture count. We leave derivation of closed-form expressions for these optima as future work – the following results are for filters tuned to their PSNR-optimal bandwidths and depth limits through exhaustive search.

5.4 Evaluation: Stanford Light Fields

Figs. 5 and 6 are typical of the output from each filter – numerical results are the peak signal-to-noise ratio (PSNR), assuming the uncorrupted input to be ideal. Fig. 5 introduces additive Gaussian noise to each light field, while Fig. 6 introduces a model of low-light camera noise, including quantization to 32 levels, intensity-dependent Poisson noise, additive Gaussian noise ($\sigma = 5\%$ maximum pixel value) and salt & pepper noise (5% density).

Visually, the hyperfan outperforms the other filters in all cases, though this will not always be true: scene elements which violate the underlying assumptions of Lambertian and non-occluding scenes will not generally conform to the hyperfan passband, and so the filter will attenuate those elements. If a scene were dominated by such elements, the filter could perform poorly – note, for example, the severely attenuated crystal ball content in Fig. 6d. These limitations are not always so jarring: the specular highlights on the Lego knights' helmets are mostly retained, for example, while the noise is mostly rejected. Furthermore, some applications can actually benefit from removal of non-Lambertian and occluding energy, for example geometric reconstruction and visual odometry.

Fig. 7 shows each method's performance for the "Lego Knights" light field over a range of aperture counts, for a variety of noise types, over a range of input noise levels, and evaluated with a range of metrics. Note that the hyperfan outperforms the others for aperture counts of five or more, and continues to improve significantly with aperture count – note the logarithmic vertical scale – confirming the scalability of the approach.

The metrics depicted in Fig. 7d are normalized to a maximum value of one. These represent the mean result over 21 levels of additive Gaussian noise with $\sigma = 10\%$ to 70% maximum pixel value. The first three metrics are, in order: PSNR, an SVD-based similarity measure,[21] and a structural similarity measure SSIM.[22] The remaining three metrics apply only to linear methods and linear noise, as they rely on separating the filter's treatment of noise and signal: by filtering the original image and the noise alone, the attenuation to each can be evaluated separately. Shown, in order, are the energy remaining when filtering the original image, the edge content of that filtered image measured as the mean magnitude of the first derivative of the image, and the inverse of the energy remaining in the filtered noise signal. Because of normalization, the best performance for all metrics is one.

Inspecting the metric results, the humble Gaussian filter takes on a prominent position in the first three metrics, even taking the lead for the SVD metric. Note, however, that the Gaussian also attenuates the most edge content. All linear methods are similar in passing signal energy, and the dual-fan outperforms the hyperfan
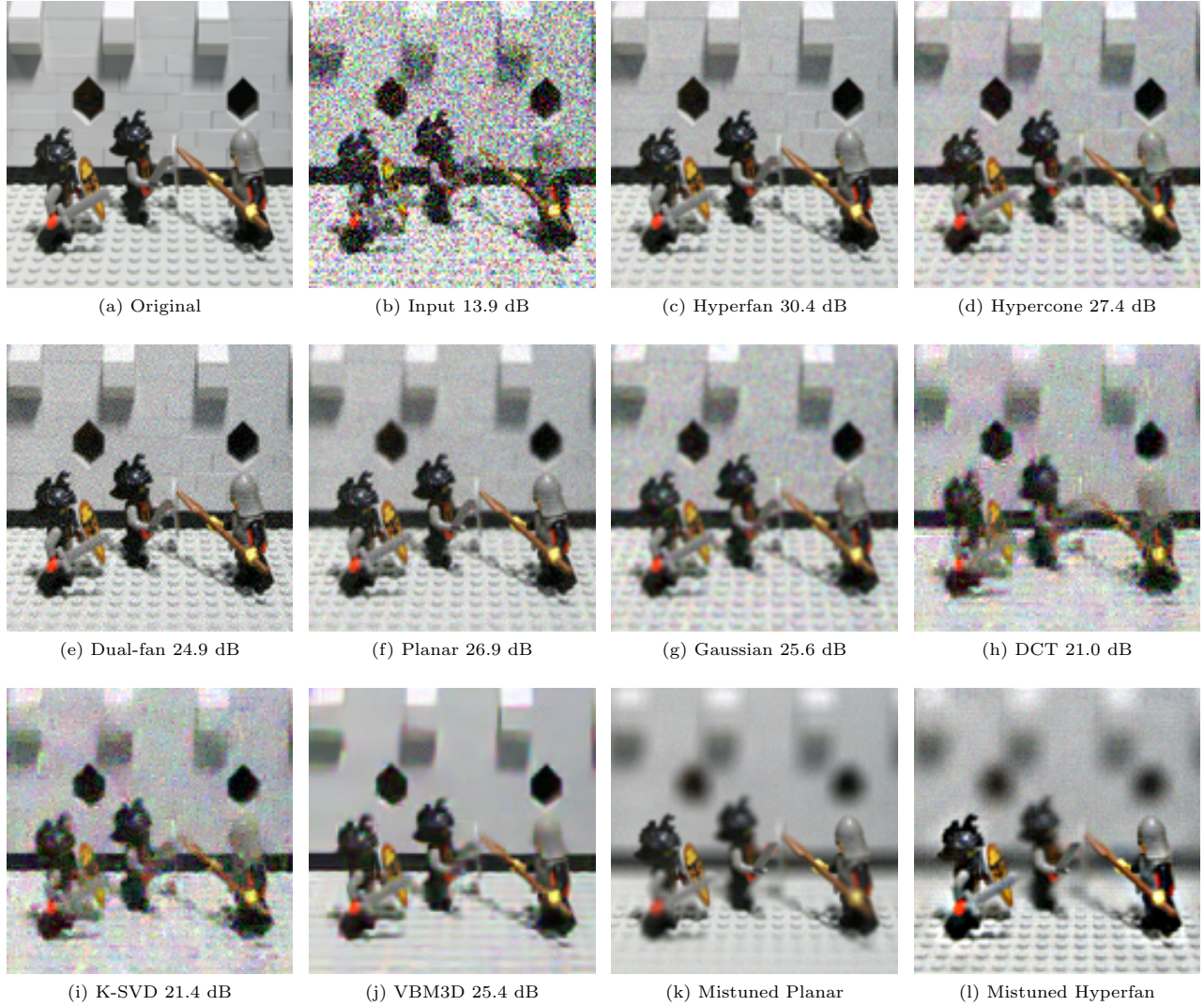
(a) Original   (b) Input 13.9 dB   (c) Hyperfan 30.4 dB   (d) Hypercone 27.4 dB

(e) Dual-fan 24.9 dB   (f) Planar 26.9 dB   (g) Gaussian 25.6 dB   (h) DCT 21.0 dB

(i) K-SVD 21.4 dB   (j) VBM3D 25.4 dB   (k) Mistuned Planar   (l) Mistuned Hyperfan

Figure 5. Filtering results for the Stanford "Lego Knights" light field: (a) the original scene, and (b) with additive white Gaussian noise; (c)–(i) show filter outputs; the depth-tunable results (c), (e) and (f) are at the PSNR-optimal balance between noise rejection and reduction in depth of field; the effects of mistuning are exaggerated in (k),(l); the hyperfan output is visually superior, with the nonlinear methods providing the most jarring artifacts, the Gaussian and planar reducing edge content, and the dual-fan and hypercone being less selective to noise
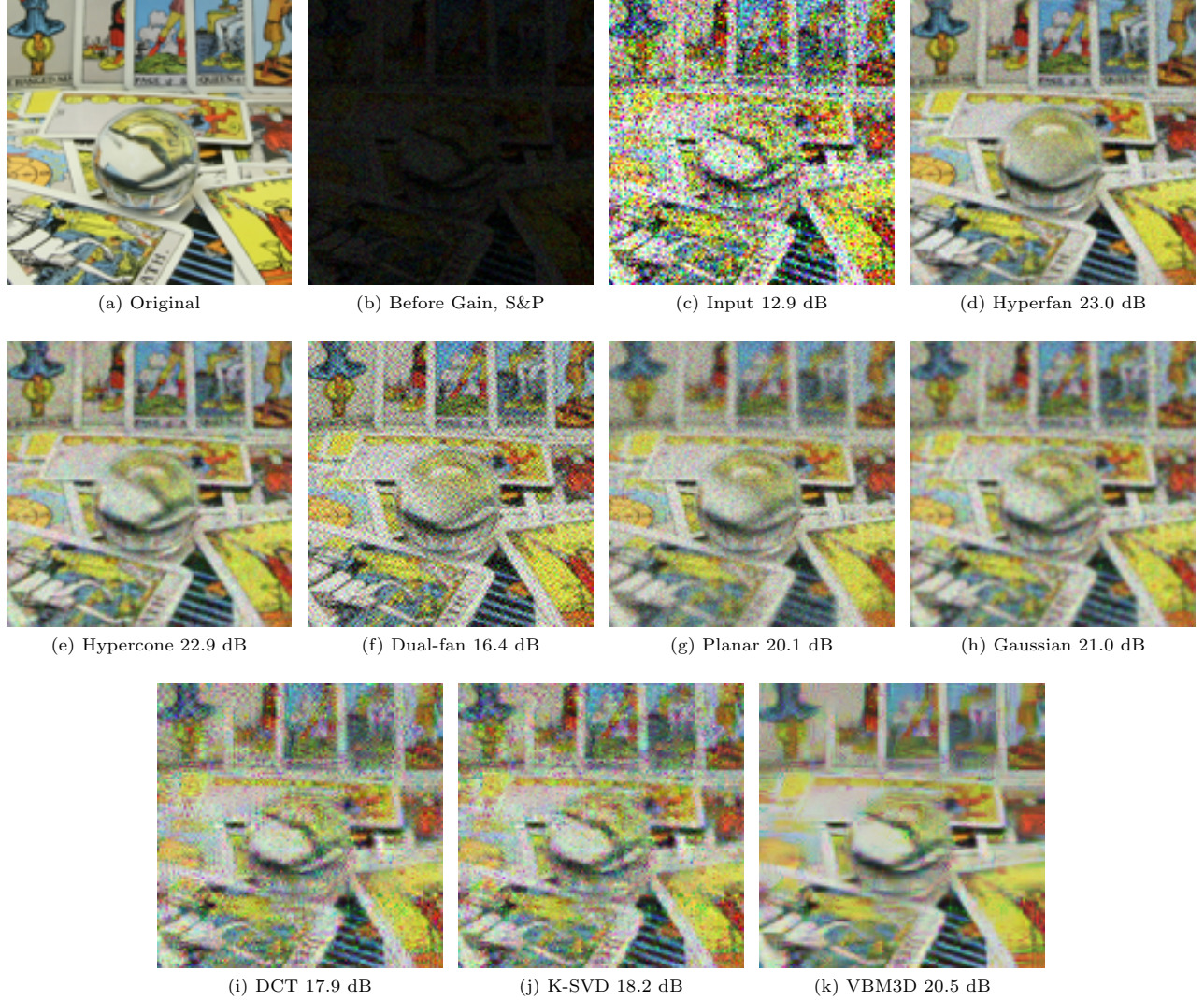
(a) Original   (b) Before Gain, S&P   (c) Input 12.9 dB   (d) Hyperfan 23.0 dB

(e) Hypercone 22.9 dB   (f) Dual-fan 16.4 dB   (g) Planar 20.1 dB   (h) Gaussian 21.0 dB

(i) DCT 17.9 dB   (j) K-SVD 18.2 dB   (k) VBM3D 20.5 dB

Figure 6. Filtering the "Tarot Coarse" light field for synthetic noise based on a camera model including quantization, Poisson, Gaussian and salt & pepper noise; (a) the original light field, (b) the low-light image prior to salt & pepper noise and gain control, (c) the gain-adjusted input including salt & pepper noise, and (d)–(k) the filter outputs; light refracting through the crystal ball violates the depth constraints, leading to attenuation of that content and a lower PSNR for depth-selective filters (d), (f), and (g); the hyperfan nevertheless arguably provides the most visually appealing result
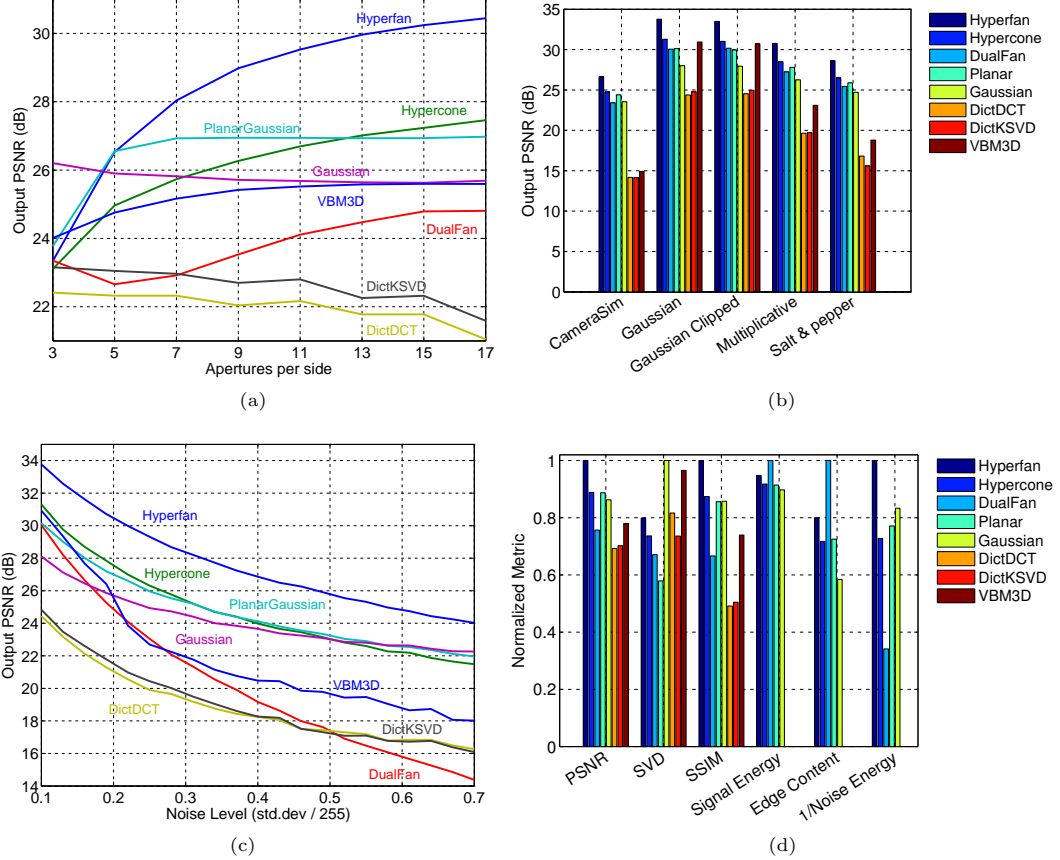
Figure 7. The performance of a variety of methods for (a) increasing aperture count, (b) a variety of noise types, (c) increasing noise level; and (d) over a variety of metrics; the hyperfan generally shows the best performance

in edge content – though it also does a poor job of attenuating noise energy, thus its weak PSNR. The nonlinear methods do well according to the SVD but a visual analysis shows that the artifacts they introduce are jarring to the human visual system. On the whole, the hyperfan attenuates the most noise energy while passing the second-to-best edge content, surpassed in this respect only by the poorly selective dual-fan. The hyperfan also dominates in structural similarity and PSNR, outperformed by its nonlinear counterparts only in the SVD metric.

Drawing on the variety of light fields available in Stanford's archive, Table 1 shows the hyperfan's performance over a range of inputs. Notable is that the output quality is high and consistent despite the varying presence of occlusions, specular reflections and refractions in the light fields – all phenomena which break the assumptions behind the filter. The weakest performance is for "Tarot Coarse", and is attributed to refraction in the scene as seen in Fig. 6d.

5.5 Evaluation: Hand-Held Plenoptic Camera Imagery

A qualitative validation was carried out on imagery collected using a Lytro consumer-grade lenselet-based hand-held light field camera, as depicted in Fig. 8. The first row depicts a low-light aquarium scene, and the second a low-light outdoor scene. Inspection of the unfiltered and filtered images shows that the hyperfan filter has significantly attenuated the noise. Note also that the specks of dirt on the side of the aquarium in the top row have been rejected by the depth selectivity of the filter – one of these is indicated by a black arrow in the inset depicting a Silver Dollar fish.

The bottom row of Fig. 8 depicts a checkerboard as imaged through turbid water. Illumination and camera were colocated, resulting in significant backscatter as seen in the leftmost image. The next image has been gain-adjusted, including removal of a low-frequency biasing term caused by backscatter. The biasing term was

Table 1. Output PSNR (dB) Over a Range of Noise Levels for the Stanford Archive

LF \ σ :	2%	5%	10%	15%	20%	25%	50%	75%	100%
Amethyst	41.91	40.35	37.98	36.10	34.58	33.50	29.81	27.52	26.09
Bracelet	40.99	38.65	35.46	33.16	31.57	30.46	26.88	24.71	23.27
Bulldozer	41.70	39.09	36.00	33.78	31.99	30.93	27.01	24.85	23.30
Bunny	45.44	41.89	38.26	35.80	34.12	32.87	28.84	26.43	24.89
Chess	44.60	41.87	38.26	35.74	34.06	32.76	28.94	26.87	25.14
Eucalyptus	43.81	41.74	38.95	37.02	35.49	34.25	30.34	28.15	26.62
JellyBeans	42.26	40.09	36.94	34.57	33.10	32.08	28.74	26.70	25.24
LegoKnights	40.48	37.16	33.96	31.83	30.32	29.03	25.56	23.47	22.20
TarotCoarse	33.82	30.73	28.12	26.36	25.14	24.34	21.42	19.77	18.60
TarotFine	40.87	37.09	34.13	32.13	30.64	29.29	25.53	23.32	21.85
Treasure	43.59	40.27	36.78	34.62	32.95	31.60	27.70	25.59	24.13
Truck	45.86	43.36	40.16	37.87	36.23	34.87	30.84	28.51	26.88
Mean	42.11	39.36	36.25	34.08	32.52	31.33	27.63	25.49	24.02
Std.Dev	3.17	3.33	3.18	3.06	2.96	2.86	2.61	2.47	2.36

estimated by low-pass filtering in the u and v dimensions. Notable is the similarity of the adjusted image to a gain-adjusted low-light image. The last two images show the output of the hyperfan filter tuned to two different depth ranges: the first is for a wide depth range including content between the camera and the checkerboard, while the final image is for a narrow filter more closely matching the geometry of this constant-depth scene. In all cases, the noise reduction effected by the hyperfan filter has clearly been significant.

5.6 Complexity and Memory

We implement the hyperfan filter in the frequency domain, making the filtering process as simple as a Fourier transform, its inverse, and a per-sample complex multiplication. Computation time for an N-sample light field is therefore constant and straightforwardly of complexity $O(N \log N)$ for the Fast Fourier Transform (FFT).

We operate on the three colour channels separately, and so the memory requirement is for a single colour channel at a time. Two buffers are required beyond the input light field buffer: the filter magnitude buffer, and a complex buffer to contain the DFT. The input light field comprises 8-bit integers, but for simplicity our implementation operates on single- or double- precision floats. For a colour light field of N samples total, our total additional memory requirement, for double-precision, is

$$M = (8 + 16)N/3 = 8N. \tag{15}$$

In practical terms, the $128 \times 128 \times 17 \times 17$ Stanford light fields shown in our results occupy about $N = 14$ MBytes, requiring an $8N/3 = 38$ MByte double-precision filter and a $16N/3 = 76$ MByte complex double-precison DFT buffer, for a total of $8N = 114$ MBytes. The single-precision implementation requires half the memory.

(a) Low-Light Input   (b) Gain-Adjusted   (c) Hyperfan-Filtered

(d) Low-Light Input   (e) Gain-Adjusted   (f) Hyperfan-Filtered

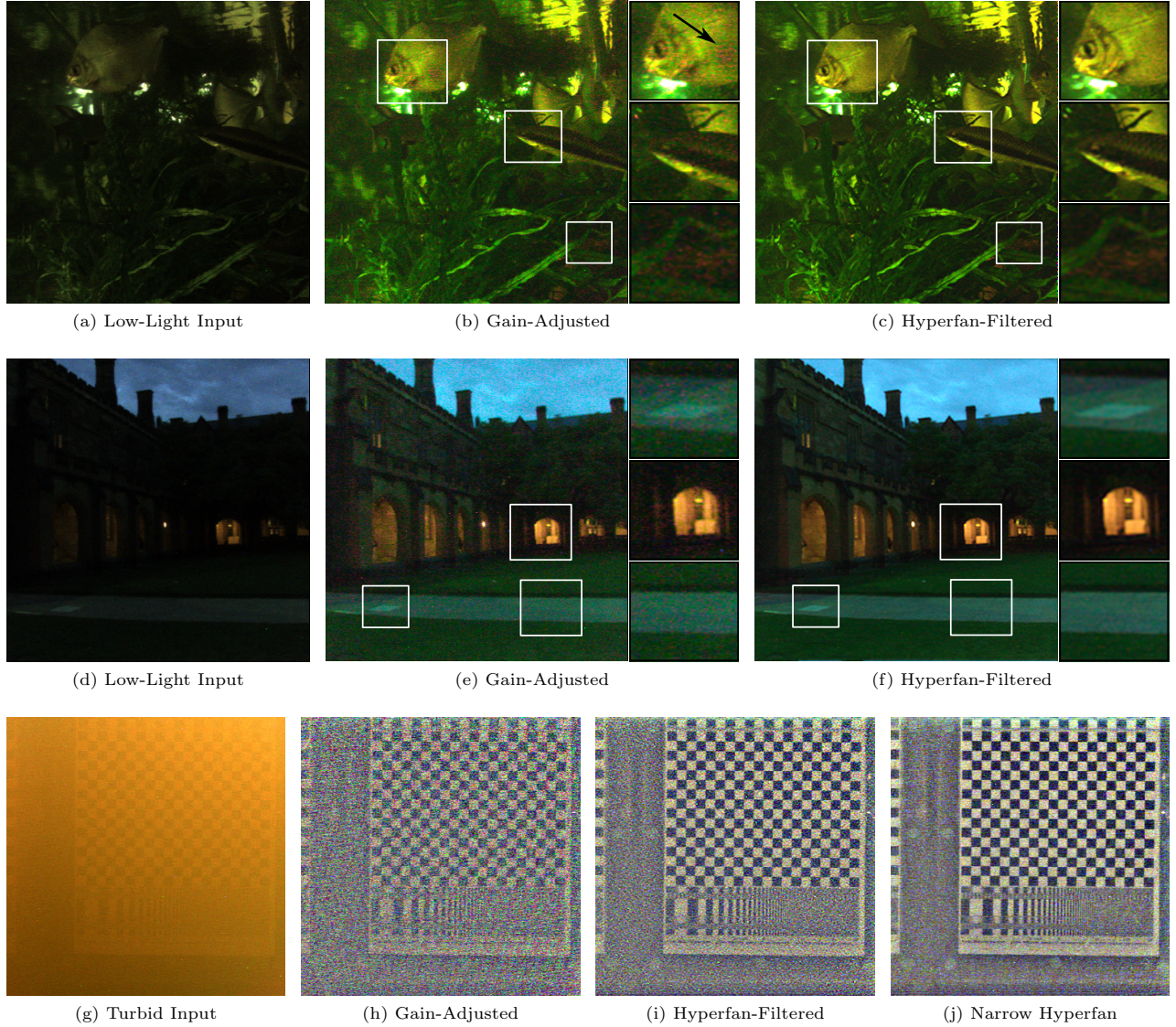(g) Turbid Input   (h) Gain-Adjusted   (i) Hyperfan-Filtered   (j) Narrow Hyperfan

Figure 8. Filtering results for a Lytro consumer-grade light field camera: (a,d,g) low-contrast input, (b,e,h) gain-adjusted images, and (c,f,i,j) filter output, showing a visible improvement in SNR. The top row demonstrates both noise rejection and depth selectivity, with specks of dirt on the side of the aquarium being attenuated based on depth – one such dirt speck is indicated by the black arrow. The bottom row demonstrates how imaging in turbid media is closely related to low-light imaging, with the only difference being that the gain-adjusted image (h) is corrected for a low-frequency biasing term prior to filtering. The final two images differ in their effective depth of field: because the scene is at a single depth, the hyperfan can be tuned to a narrow depth of field to yield more dramatic noise rejection (j).

6. CONCLUSIONS AND FUTURE WORK

We have established that the frequency-domain region of support (ROS) of a light field image is a hyperfan at the intersection of a dual-fan and a hypercone. We have designed, implemented and tested a novel filter which selectively passes this ROS. This approach to light field denoising is linear and featureless, operating efficiently and in constant time independent of scene complexity.

We have demonstrated the filter outperforming a range of linear and nonlinear alternatives over a range of conditions including noise type, noise level, aperture count and scene content. Test scenes included examples of occlusion and non-Lambertian surfaces. Results were shown for twelve light fields from the Stanford Light Field Archive, including representative images and quantitative results over a range of metrics. The filter was shown to be effective at removing noise in all cases, generally outperforming the other methods we evaluated including planar, dual-fan, overcomplete DCT, K-SVD and video-based VBR3D methods. We also showed that the hyperfan filter's performance scales with aperture count. Finally, we demonstrated the filter on imagery collected with the Lytro consumer-grade light field camera, including scenes with low-light and turbid (murky) water. In all cases the hyperfan filter was demonstrated to visibly improve image quality.

We see the frequency-domain implementation of the hyperfan as a first step, leaving for future work implementation in the spatial domain and with limited-precision arithmetic, allowing efficient deployment of the technique in a range of scenarios including embedded applications. Automated means of selecting filter parameters would be desirable. Finally, we believe the hyperfan filter could be useful for a range of interesting tasks, including compression and interpolation for rendering.

Acknowledgment

REFERENCES

[1] T. Bishop and P. Favaro, "The light field camera: Extended depth of field, aliasing, and superresolution," *Pattern Analysis and Machine Intelligence, IEEE Trans. on* **34**, pp. 972–986, May 2012.

[2] A. Veeraraghavan, R. Raskar, A. Agrawal, A. Mohan, and J. Tumblin, "Dappled photography: Mask enhanced cameras for heterodyned light fields and coded aperture refocusing," *ACM Trans. on Graphics* **26**(3), p. 69, 2007.

[3] R. Ng, M. Levoy, M. Brédif, G. Duval, M. Horowitz, and P. Hanrahan, "Light field photography with a hand-held plenoptic camera," *Computer Science Technical Report CSTR* **2**, 2005.

[4] R. Ng, "Fourier slice photography," in *ACM Trans. on Graphics (TOG)*, **24**, pp. 735–744, ACM, Jul 2005.

[5] B. Wilburn, N. Joshi, V. Vaish, E. Talvala, E. Antunez, A. Barth, A. Adams, M. Horowitz, and M. Levoy, "High performance imaging using large camera arrays," *ACM Trans. on Graphics (TOG)* **24**(3), pp. 765–776, 2005.

[6] A. Buades, B. Coll, J. Morel, *et al.*, "A review of image denoising algorithms, with a new one," *SIAM Journal on Multiscale Modeling and Simulation* **4**(2), pp. 490–530, 2005.

[7] O. Guleryuz, "Weighted averaging for denoising with overcomplete dictionaries," *Image Processing, IEEE Trans. on* **16**(12), pp. 3020–3034, 2007.

[8] M. Aharon, M. Elad, and A. Bruckstein, "K-SVD: An algorithm for designing overcomplete dictionaries for sparse representation," *Signal Processing, IEEE Trans. on* **54**(11), pp. 4311–4322, 2006.

[9] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *Image Processing, IEEE Trans. on* **15**(12), pp. 3736–3745, 2006.

[10] K. Dabov, A. Foi, and K. Egiazarian, "Video denoising by sparse 3D transform-domain collaborative filtering," in *Proc. 15th European Signal Processing Conf.*, p. 7, 2007.

[11] M. Levoy, B. Chen, V. Vaish, M. Horowitz, I. McDowall, and M. Bolas, "Synthetic aperture confocal imaging," *ACM Trans. on Graphics (TOG)* **23**(3), pp. 825–834, 2004.

[12] D. G. Dansereau and L. T. Bruton, "A 4-D Dual-Fan Filter Bank for Depth Filtering in Light Fields," *IEEE Trans. on Signal Processing* **55**(2), pp. 542–549, 2007.

[13] D. G. Dansereau and S. B. Williams, "Seabed modeling and distractor extraction for mobile AUVs using light field filtering," in *Robotics and Automation (ICRA), IEEE Intl. Conference on*, pp. 1634–1639, IEEE, May 2011.

[14] H. Yang, M. Pollefeys, G. Welch, J. Frahm, and A. Ilie, "Differential camera tracking through linearizing the local appearance manifold," in *Computer Vision and Pattern Recognition (CVPR), IEEE Conference on*, pp. 1–8, IEEE, 2007.

[15] D. G. Dansereau, I. Mahon, O. Pizarro, and S. B. Williams, "Plenoptic Flow: Closed-Form Visual Odometry for Light Field Cameras," in *Intelligent Robots and Systems (IROS), IEEE/RSJ Intl. Conf. on*, pp. 4455–4462, IEEE, Sept 2011.

[16] M. Levoy and P. Hanrahan, "Light field rendering," in *Proceedings of the 23rd annual conference on Computer graphics and interactive techniques*, pp. 31–42, ACM, 1996.

[17] R. Bolles, H. Baker, and D. Marimont, "Epipolar-plane image analysis: An approach to determining structure from motion," *International Journal of Computer Vision* **1**(1), pp. 7–55, 1987.

[18] D. G. Dansereau and L. Bruton, "A 4D frequency-planar IIR filter and its application to light field processing," in *Proceedings of the Intl. Symposium on Circuits and Systems*, **4**, pp. 476–479, May 2003.

[19] A. Levin, S. Hasinoff, P. Green, F. Durand, and W. Freeman, "4D frequency analysis of computational cameras for depth of field extension," *ACM Trans. on Graphics (TOG)* **28**(3), p. 97, 2009.

[20] R. Ansari, "Efficient IIR and FIR fan filters," *IEEE Trans. on circuits and systems* **34**(8), pp. 941–945, 1987.

[21] A. Shnayderman, A. Gusev, and A. Eskicioglu, "An SVD-based grayscale image quality measure for local and global assessment," *Image Processing, IEEE Trans. on* **15**(2), pp. 422–429, 2006.

[22] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *Image Processing, IEEE Trans. on* **13**(4), pp. 600–612, 2004.